Measuring subluxation of the hemiplegic shoulder: Reliability of a method

Snels, I.A.K.; Beckerman, H.; ten Kate, J.J.; Lankhorst, G.J.; Bouter, L.M.

published in
Neurorehabilitation and Neural Repair
2001

DOI (link to publisher)
10.1177/154596830101500314

document version
Publisher's PDF, also known as Version of record

Neurorehabilitation and Neural Repair 2001;15:249–254

Brief Communication

Measuring Subluxation of the Hemiplegic Shoulder: Reliability of a Method

*†Ingrid A.K. Snels, *†Heleen Beckerman, *Joke J. ten Kate, *†Gustaaf J. Lankhorst, and †Lex M. Bouter

Objective: Subluxation of the shoulder after stroke can be measured according to the method described by Van Langenberghe and Hogan. **Methods:** To evaluate the reliability of this method, the shoulder radiographs of 25 patients were available for this study. Two independent raters each assessed these radiographs twice. **Results:** The intrarater reliability was good: percentage of agreement was 88 and 84%, weighted κ, 0.69 [95% confidence interval (CI), 0.38–1.0] and 0.78 (95% CI, 0.60–0.95) for raters 1 and 2, respectively. The interrater reliability was poor: percentage of agreement was 36 and 28%, κ, 0.11 (95% CI, 0.0–0.31) and 0.09 (95% CI, 0.0–0.23) in sessions 1 and 2, respectively. Subsequently the original method was adjusted by combining two categories (no subluxation and beginning subluxation) into one ("no clinically important subluxation"). **Conclusions:** After this adjustment of the categories, the interrater reliability improved [percentage of agreement, 72%, and κ, 0.49 (95% CI, 0.18–0.80)], but did not reach acceptable values. **Key Words:** Hemiplegic shoulder pain—Subluxation—Measurement.

Subluxation of the shoulder is a common complication after stroke, and is often mentioned as a possible cause for hemiplegic shoulder pain (1–7). Subluxation can be measured clinically or by taking a radiograph of the shoulder. Clinical methods are palpation of the subacromial gap (1, 6, 8–15), or measuring this gap with a plexiglass jig (8, 11, 16) or a calliper (8). A third clinical method for diagnosing subluxation is to measure the difference in arm length with a calliper (11, 13). A radiograph for the evaluation of subluxation should be made with the arm in a dependent (unsupported) position (17 18). Most frequently, an anteroposterior radiograph of one or both shoulders is used (7, 9, 11, 15, 17, 19–25). In addition to anteroposterior radiographs, radiographs in other planes are sometimes used (e.g., 30 degrees, 45 degrees, in plane of scapula) (5, 8, 13, 15, 23, 26, 27). Kobayashi et al. (28) took an anteroposterior radiograph of both shoulders with a stress test to diagnose subluxation.

In addition to descriptions of the technique of taking a radiograph of the hemiplegic shoulder, several methods have been described to evaluate subluxation from the resulting radiographs. In broad outline, there are two ways to evaluate these radiographs. First, subluxation is defined as the distance between the inferior part of the acromion and the superior part of the humerus (7, 11, 15, 20). Second, subluxation is defined in terms of the position of the humerus in relation to the glenoid fossa (9, 21, 22, 25, 28–31).

From the *Department of Rehabilitation Medicine, University Hospital, and †Institute for Research in Extramural Medicine, Vrije Universiteit, Amsterdam, The Netherlands.

Address correspondence and reprint requests to Heleen Beckerman, Department of Rehabilitation Medicine, University Hospital, Vrije Universiteit, PO Box 7057, 1007 MB Amsterdam, The Netherlands. E-mail h.beckerman@vumc.nl

A randomized clinical trial (RCT) was carried out to determine the effectiveness of triamcinolone acetonide injections for hemiplegic shoulder pain (32). As subluxation could be a prognostic factor, it was measured before the start of the intervention, according to the method described by Van Langenberghe and Hogan (29). Van Langenberghe and Hogan assessed the degree of subluxation of 48 stroke patients with shoulder pain (29). Their method was chosen because it seemed easy to apply, the reported intra- and interreliability are very good, and it can be applied to (routine) anteroposterior radiographs. However, Van Langenberghe and Hogan (29) calculated only percentages of agreement and excluded some radiographs. This article reports on a reevaluation of the intra- and interrater reliability of their method in our series of patients.

Methods

Patients and Procedures

Patients with hemiplegic shoulder pain after stroke participated in the RCT if they fulfilled the following criteria: ⩾2 weeks of pain ⩾4.0 on a visual analog scale of 0–10 in their hemiplegic shoulder, and no history of shoulder problems before stroke. At baseline, a radiograph of the hemiplegic shoulder was made in anteroposterior position with an unsupported, dependent arm. These radiographs were rated according the 5-point categorization described by Van Langenberghe and Hogan (29): 0, normal (no subluxation): the whole curvature of the glenoid fossa is opposed by and parallel to the humeral head; 1, V-shaped widening (beginning subluxation): the whole curvature of the glenoid fossa is opposed by the humeral head with loss of parallelism; 2, moderate subluxation: there is inferior subluxation of the humeral head, but its most superior margin is above the line perpendicularly bisecting the line connecting the most superior and the most inferior margins of the glenoid fossa; 3, severe subluxation: as 2, but the superior margin of the humeral head is not above the bisecting line; and 4, dislocation: the most superior margin of the humeral head is not above the most inferior margin of the glenoid fossa. If it was not possible to allocate a radiograph to one of these five categories, it was rated as "not assessable."

Raters and Rating

Before rating took place, the patient identity was blinded on the radiographs, and in each session, the radiographs were rated in a different random order. Two experienced rehabilitation physicians (G.J.L., J.J.tK.) independently rated the radiographs twice, with a time interval of ⩾4 weeks. Subsequently, a consensus meeting was held to establish one score for each patient. Consensus was necessary to investigate the possible relation between (degree of) subluxation and pain. During this consensus meeting, it became clear that the distinction between category 0 (no subluxation) and category 1 (beginning subluxation) was most difficult. Therefore, to improve clinical usefulness, these two categories were combined into one category ("no clinically important subluxation"). After this adjustment, the same rehabilitation physicians independently rated the radiographs again, using the modified scoring chart.

Analysis

For both the intrarater and interrater reliability, percentages agreement and weighted κ values were calculated. The interrater reliability was assessed for all three rating sessions (33, 34). For the interpretation of the κ values, the criteria proposed by Altman (33) were used: κ values <0.20 may be taken to represent poor agreement; values between 0.21 and 0.40, fair; 0.41 and 0.60, moderate; 0.61 and 0.80, good; and values >0.80, very good agreement. SPSS 9.0 was used for the analysis.

Results

In total, the radiographs of 25 patients were available for evaluation of the degree of subluxation. Table 1

Table 1. *Characteristics of 25 patients with hemiplegic shoulder pain after stroke*

	No subluxation[a] (n = 16)	Subluxation[a] (n = 5)	Not Assessable[a] (n = 4)
Median age (yr)	62	54	56.5
(range)	(48–80)	(44–65)	(51–70)
Gender (m/f)	9/7	1/4	3/1
Median pain (0–10)	5.2	2.6	4.9
(range)	(1.3–10)	(2.4–6.7)	(4–5.8)
Time since onset of stroke			
<6 mo	6	5	2
⩾6 mo	10	0	2

[a]Outcome of the consensus meeting.

Table 2A. *Intrarater reliability of rater 1*

	Session 2				
	Normal	V-shaped	Moderate subluxation	Not assessable	Total
Session 1					
Normal	15	0	0	0	15
V-shaped	1	2	0	0	3
Moderate subluxation	0	0	5	0	5
Not assessable	2	0	0	0	2
Total	18	2	5	0	25

Percentage of agreement, 88%.
κ = 0.69 (95% CI, 0.38–1.0).

Table 2B. *Intrarater reliability of rater 2*

	Session 2				
	Normal	V-shaped	Moderate subluxation	Not assessable	Total
Session 1					
Normal	4	2	0	0	6
V-shaped	0	1	0	1	2
Moderate subluxation	0	0	2	0	2
Not assessable	1	0	0	14	15
Total	5	3	2	15	25

Percentage of agreement, 84%.
κ = 0.78 (95% CI, 0.60–0.95).

presents some characteristics of these patients. None of the patients had severe subluxation (category 3) or dislocation (category 4). The percentages of agreement were 88 and 84% for rater 1 and rater 2, respectively. The weighted κ for the intrarater reliability was 0.69 (95% CI, 0.38–1.0) and 0.78 (95% CI, 0.60–0.95) for rater 1 and rater 2, respectively (Table 2A and B). With the unadjusted five categories, the percentages of agreement between both raters were 36 and 28% in sessions 1 and 2, respectively; the interrater reliability was 0.11 (95% CI, 0.0–0.31) and 0.09 (95% CI, 0.0–0.23) for the two sessions (Table 3A and B). After adjusting the categories, the percentage of agreement between both raters was 72%, and the interrater reliability increased to 0.49 (95% CI, 0.18–0.80; Table 4).

Discussion

The results of our reliability study are less favorable than those of Van Langenberghe and Hogan, who reported excellent interrater (100%) and intrarater reliability (92%) (29). Two factors probably contributed to the results of Van Langenberghe and Hogan. In the first place, they excluded four radiographs from the analysis because these were consistently rated differently by the two assessors. In the second place, they calculated these reliabilities as the agreement between the first and the second rating, and the two ratings made by each assessor. Percentages of agreement, however, do not take into account the probability of some chance agreement. κ does take this into account, but no κ values can be calculated from their published data. Calculations of weighted κ values in the present study showed good intrarater reliability, but poor interrater reliability (33). During the consensus meeting, it became clear that the distinction between no subluxation (category 0) and V-shaped widening (category 1) was most difficult to make for both raters. Rater 2 often rated these cases as "not assessable"; rater 1 categorized the radiographs into one of the two categories. After the decision to combine category 0 and category 1 ("no clinically important subluxation"), the interrater reliability improved, but a weighted κ of 0.49 (95% CI,

Table 3A. *Interrater reliability during session 1*

	Rater 1				
	Normal	V-shaped	Moderate subluxation	Not assessable	Total
Rater 2					
Normal	4	2	0	0	6
V-shaped	0	1	1	0	2
Moderate subluxation	0	0	2	0	2
Not assessable	11	0	2	2	15
Total	15	3	5	2	25

Percentage of agreement, 36%.
κ = 0.11 (95% CI, 0.0–0.31).

Table 3B. *Interrater reliability during session 2*

	Rater 1				
	Normal	V-shaped	Moderate subluxation	Not assessable	Total
Rater 2					
Normal	5	0	0	0	5
V-shaped	1	1	1	0	3
Moderate subluxation	0	0	2	0	2
Not assessable	12	1	2	0	15
Total	18	2	5	0	25

Percentage of agreement, 28%.
κ = 0.09 (95% CI, 0.0–0.23).

0.18–0.80) is only moderate, and not good (33). Moreover, in reducing the number of categories, κ will always increase (33). The distribution of the degrees of subluxation over the different categories was unequal, and more equal distribution might have resulted in a higher κ (33). Roy et al. (24) made the same adjustment (combining categories 0 and 1), but reported no results concerning reliability. No other studies reported any difficulties with the discrimination between categories (35, 36). It is to be noted that our study did not test the reliability of the whole scale, but only part of it, because there were no cases with severe subluxation or dislocation in the sample. Another limitation is the fact that our study included only two raters.

Table 4. *Interrater reliability during third session (after adaptation of the categories)*

	Rater 2			
	Normal orV-shaped	Moderate subluxation	Not assessable	Total
Rater 1				
Normal or V-shaped	12	1	2	15
Moderate subluxation	1	4	2	7
Not assessable	1	0	2	3
Total	14	5	6	25

Percentage of agreement, 72%.
κ = 0.49 (95% CI, 0.18–0.80).

Several explanations are possible for the disappointingly low interrater reliability. First, the raters were experienced rehabilitation physicians, not radiologists, but with clearly described categories, every experienced physician is able to categorize routine shoulder radiographs. Therefore, it is not likely that this had any great influence on the results. Second, the two raters were not familiar with the method before they applied it. This could have influenced the results, but even in the third session after an intensive consensus meeting, the interrater reliability was only moderate. Third, the radiographs took place in different institutions, and the diagnostic radiology technicians who made the radiographs had received written instructions to make an anteroposterior radiograph with an unsupported, dependent arm. Perhaps this was insufficient to guarantee good and equal quality of the radiographs. According to the rating rehabilitation physicians, several radiographs were of suboptimal quality. However, this is often the case in normal clinical practice.

In summary, this study found the intrarater reliability of the method described by Van Langenberghe and Hogan to be good, and the interrater reliability poor to moderate. Adaptation of the assessment method was partly successful: weighted κ increased, but did not reach acceptable values. We recommend that clinicians, wanting to quantify subluxation, take care of optimization of radiographic techniques and hold consensus meetings with colleagues to become aware of differences in judgment and to improve assessment.

References

1. Bohannon RW, Andrews AW. Shoulder subluxation and pain in stroke patients. *Am J Occup Ther* 1990;44:507–9.
2. Bruton JD. Shoulder pain in stroke patients with hemiplegia or hemiparesis following a cerebrovascular accident. *Physiotherapy* 1985;71:2–4.
3. Eliasar Bouwman M, Gorter S, Van Dolder R. De relatie tussen subluxatie en de pijnlijke hemiplegische schouder: Een literatuurstudie. *Ned Tijdschr Fysiother* 1994;104:170–4.
4. Griffin JW. Hemiplegic shoulder pain. *Phys Ther* 1986;66:1884–93.
5. Ikai T, Tei K, Yoshida K, et al. Evaluation and treatment of shoulder subluxation in hemiplegia: relationship between subluxation and pain. *Am J Phys Med* 1998;77:421–6.
6. Wanklyn P, Forster A, Young J. Hemiplegic shoulder pain (HSP): natural history and investigation of associated features. *Disabil Rehabil* 1996;18:497–501.
7. Zorowitz RD, Hughes MB, Idank D, et al. Shoulder pain and subluxation after stroke: correlation or coincidence? *Am J Occup Ther* 1996;50:194–201.
8. Boyd EA, Goudreau L, O'Riain MD, et al. A radiological measure of shoulder subluxation in hemiplegia: its reliability and validity. *Arch Phys Med Rehabil* 1993;74:188–93.
9. Chantraine A, Baribeault A, Uebelhart D, et al. Shoulder pain and dysfunction in hemiplegia: effects of functional electrical stimulation. *Arch Phys Med Rehabil* 1999;80:328–31.
10. Chino N. Electrophysiological investigation on shoulder subluxation in hemiplegics. *Scand J Rehabil Med* 1981;13:17–21.
11. Hall J, Dudgeon B, Guthrie M. Validity of clinical measures of shoulder subluxation in adults with poststroke hemiplegia. *Am J Occup Ther* 1995;49:526–33.
12. Lee KH, Khunadorn F. Painful shoulder in hemiplegic patients: a study of the suprascapular nerve. *Arch Phys Med Rehabil* 1986;67: 818–20.
13. Prevost R, Arsenault AB, Dutil E, et al. Shoulder subluxation in hemiplegia: a radiologic correlational study. *Arch Phys Med Rehabil* 1987;68:782–5.
14. Tobis JS. Problems in rehabilitation of the hemiplegic patient. *N Y State J Med* 1957;57:1377–80.
15. Williams R, Taffs L, Minuk T. Evaluation of two support methods for the subluxated shoulder of hemiplegic patients [published erratum appeared in Phys Ther 1988;68:1969]. *Phys Ther* 1988;68: 1209–14.
16. Hayes KW, Sullivan JE. Reliability of a new device used to measure shoulder subluxation. *Phys Ther* 1989;69:762–7.
17. Fitzgerald-Finch OP, Gibson II. Subluxation of the shoulder in hemiplegia. *Age Ageing* 1975;4:16–8.
18. Moskowitz H, Goodman CR, Smith E, et al. Hemiplegic shoulder. *N Y State J Med* 1969;69:548–50.
19. Baker LL, Parker K. Neuromuscular electrical stimulation of the muscles surrounding the shoulder. *Phys Ther* 1986;66:1930–7.
20. Brooke MM, De Lateur BJ, Diana-Rigby GC, et al. Shoulder subluxation in hemiplegia: effects of three different supports. *Arch Phys Med Rehabil* 1991;72:582–6.
21. Najenson T, Yacubovich E, Pikielni SS. Rotator cuff injury in shoulder joints of hemiplegic patients. *Scand J Rehabil Med* 1971;3:131–7.
22. Ring H, Leillen B, Server S, et al. Temporal changes in electrophysiological, clinical and radiological parameters in the hemiplegic's shoulder. *Scand J Rehabil Med Suppl* 1985;12:124–7.
23. Rizk TE, Christopher RP, Pinals RS, et al. Arthrographic studies in painful hemiplegic shoulders. *Arch Phys Med Rehabil* 1984;65: 254–6.
24. Roy CW, Sands MR, Hill LD. Shoulder pain in acutely admitted hemiplegics. *Clin Rehabil* 1994;8:334–40.
25. Shai G, Ring H, Costeff H, et al. Glenohumeral malalignment in the hemiplegic shoulder: an early radiologic sign. *Scand J Rehabil Med* 1984;16:133–6.
26. Carpenter GI, Millard PH. Shoulder subluxation in elderly inpatients. *J Am Geriatr Soc* 1982;30:441–6.
27. Culham EG, Noce RR, Bagg SD. Shoulder complex position and glenohumeral subluxation in hemiplegia. *Arch Phys Med Rehabil* 1995;76:857–64.
28. Kobayashi H, Onishi H, Ihashi K, et al. Reduction in subluxation and improved muscle function of the hemiplegic shoulder joint after therapeutic electrical stimulation. *J Electromyogr Kinesiol* 1999;9:327–36.
29. Van Langenberghe HV, Hogan BM. Degree of pain and grade of subluxation in the painful hemiplegic shoulder. *Scand J Rehabil Med* 1988;20:161–6.
30. Prevost R, Arsenault AB, Dutil E, et al. Rotation of the scapula and shoulder subluxation in hemiplegia. *Arch Phys Med Rehabil* 1987;68:786–90.
31. De Bats M, De Bisschop G, Bardot A, et al. La subluxation inférieure de l'épaule chez l'hémiplégique. *Ann Med Phys* 1974;2: 185–213.
32. Snels IAK, Beckerman H, Twisk JWR, et al. Effect of triamci-

nolone acetonide injections on hemiplegic shoulder pain: a randomized clinical trial. *Stroke* 2000;31:2396–401.
33. Altman DG. *Practical statistics for medical research*. 2nd ed. London: Chapman & Hall, 1991.
34. Fleiss JL. *Statistical methods for rates and proportions*. 2nd ed. New York: John Wiley & Sons, 1981.
35. Feys HM, De Weerdt WJ, Selz BE, et al. Effect of a therapeutic intervention for the hemiplegic upper limb in the acute phase after stroke: a single-blind, randomized, controlled multicenter trial. *Stroke* 1998;29:785–92.
36. Linn SL, Granat MH, Lees KR. Prevention of shoulder subluxation after stroke with electrical stimulation. *Stroke* 1999;30:963–8.